

Review

Machine Learning and Deep Learning in Cardiothoracic Imaging: A Scoping Review

Bardia Khosravi ^{1,2,†}, Pouria Rouzrok ^{1,2,†}, Shahriar Faghani ¹, Mana Moassefi ¹, Sanaz Vahdati ¹, Elham Mahmoudi ¹, Hamid Chalian ³ and Bradley J. Erickson ^{1,*}

¹ Radiology Informatics Lab (RIL), Department of Radiology, Mayo Clinic, Rochester, MN 55905, USA

² Orthopedic Surgery Artificial Intelligence Laboratory (OSAIL), Department of Orthopedic Surgery, Mayo Clinic, Rochester, MN 55905, USA

³ Department of Radiology, Cardiothoracic Imaging, University of Washington, Seattle, WA 98195, USA

* Correspondence: bje@mayo.edu; Tel.: +1-(507)-284-8548

† These authors contributed equally to this work.

Abstract: Machine-learning (ML) and deep-learning (DL) algorithms are part of a group of modeling algorithms that grasp the hidden patterns in data based on a training process, enabling them to extract complex information from the input data. In the past decade, these algorithms have been increasingly used for image processing, specifically in the medical domain. Cardiothoracic imaging is one of the early adopters of ML/DL research, and the COVID-19 pandemic resulted in more research focus on the feasibility and applications of ML/DL in cardiothoracic imaging. In this scoping review, we systematically searched available peer-reviewed medical literature on cardiothoracic imaging and quantitatively extracted key data elements in order to get a big picture of how ML/DL have been used in the rapidly evolving cardiothoracic imaging field. During this report, we provide insights on different applications of ML/DL and some nuances pertaining to this specific field of research. Finally, we provide general suggestions on how researchers can make their research more than just a proof-of-concept and move toward clinical adoption.



Citation: Khosravi, B.; Rouzrok, P.; Faghani, S.; Moassefi, M.; Vahdati, S.; Mahmoudi, E.; Chalian, H.; Erickson, B.J. Machine Learning and Deep Learning in Cardiothoracic Imaging: A Scoping Review. *Diagnostics* **2022**, *12*, 2512. <https://doi.org/10.3390/diagnostics12102512>

Academic Editor: Sameer Antani

Received: 17 September 2022

Accepted: 15 October 2022

Published: 17 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Artificial intelligence (AI) is a broad term used to define systems that perform tasks that typically require human intelligence [1]. Early efforts to create such systems in medicine focused on using known relationships, casualties, and decision logic to develop an “intelligent” algorithm, much like diagnostic flowcharts but executed by a computer. In the 21st century, with the widespread adoption of electronic health records (EHR), “big data” came into existence [2]. Big data refers to massive multidimensional collections of information about each individual, for example, their demographic information, lab test results, medication history, and imaging studies, just to name a few. Using traditional statistical methods to interpret this data is not optimal, as these methods cannot capture the complexities of multidimensional data [3]. Therefore, researchers turned to machine-learning (ML) algorithms to harness this information.

ML is a subfield of AI that encompasses algorithms that “learn” the relationships between data elements by seeing many examples [4]. This makes ML algorithms data-driven, meaning that the input data’s quality (and quantity) determines their performance. Deep learning (DL) is a group of ML algorithms that extract relationships in the data through multilayer neural networks, resembling a human cognition system [5]. Due to their hierarchical architecture, DL algorithms have the capacity to abstract complex information from images. They have proven their potential by outperforming humans in several natural-image tasks, such as handwritten image classification and face detection [6–8].

Medical imaging is the perfect place to harness the power of these ML and DL algorithms. As an example, just in 2016, there were more than 75 million computed tomography (CT) examinations performed in the United States [9]. Considering this number and the vast information in each study, the potential to use ML algorithms to extract information to help diagnose different diseases is immense. Additionally, ML and DL can be used to extract parameters from the image to predict the prognosis or survival of a patient or optimize clinical workflows [10]. There are other potential applications for DL in image reconstruction, which are listed in the applications section in the results of this report.

Developing these algorithms requires some programming background, but, in recent years, there have been some no-code solutions, called *AutoML*. These allow researchers to focus on the input data, and the rest, including model development and fine-tuning, is taken care of by these *AutoML* softwares. These produce results comparable to manually developed models and have lowered the bar to use ML/DL algorithms by clinical researchers [11].

Several studies have explored radiologists' attitudes toward the progression of AI in their field and their knowledge about it [12]. In a survey by Huisman et al. of more than 1000 participants, 69% reported some knowledge of AI, with 11% denoting their knowledge as advanced [13]. In contrast, Ooi et al. reported that 65% of their participants evaluated their knowledge of AI as novice at best [14]. Despite these gaps in knowledge, radiologists have a consensus that AI will dramatically impact their field in ten years or even sooner [15–17]. Additionally, in several reports, radiologists have expressed that they do not believe AI will replace them in the foreseeable future [14,18,19]. The model usually proposed is a cooperation of AI and radiologists, with AI facilitating and increasing the efficiency of radiologists and improving the diagnostic and prognostic workflows [20]. A radiologist might not need to know the technical details of ML/DL algorithms; however, knowing some general concepts would help one prepare for an even-more technology-rich practice.

Cardiothoracic imaging is one of the early adopters of ML/DL algorithms. This is mainly attributed to the public release of large datasets early in the days that DL was gaining attraction in the medical field [21]. These datasets provided a valuable resource for researchers and developers outside the medical field to test and benchmark their algorithms. Additionally, with the COVID-19 pandemic, an arsenal of researchers have tried to apply computer-aided diagnostic tools to help with the early diagnosis of the disease and predict the prognosis of patients [22]. With the variety of modalities used in cardiothoracic imaging, there is a potential to help augment radiologists' abilities in their routine clinical practice.

In this report, we review papers on cardiothoracic image analysis using AI and try to paint a big picture of advances in this field. For the presented results, we highlight some nuances of applying AI to medical problems, specifically cardiothoracic imaging. Finally, we summarize the findings by providing some useful tips for radiologists interested in performing applied AI research.

2. Materials and Methods

In this review, we addressed *how ML/DL are used in cardiothoracic imaging*. Inclusion criteria included (a) peer-reviewed articles on (b) cardiothoracic imaging, with (c) English full-text, that (d) have used ML or DL algorithms in their study. We excluded articles (a) focusing on non-imaging studies, specifically pathology images, and (b) all non-original research papers, including review articles or systematic analyses. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) extension for scoping reviews [23].

To identify relevant studies, a systematic search was done in the MedLine database. The search included three parts, one to include organs that are located in the thoracic area, another to target the ML/DL component of their study, and, finally, a part to search for relevant modalities. Supplementary Table S1 shows the terms used in this phase. Additionally, due to the effect of COVID-19 on the number of publications both in the cardiothoracic

imaging field in general and also AI research, we included the term “COVID-19” to encompass the latest advancements in this field [24]. The search was confined to the period 1 January 2012 to 31 May 2022 to be inclusive of the most recent advances in the field. The beginning timepoint was selected as the first scalable convolutional neural network (i.e., a subclass of DL models that work on image analysis) introduced back in 2012 and significantly improved the state-of-the-art results in natural-image classification [25]. The search was initially carried out on 15 July 2022 and later updated on 27 August 2022.

For each study, 11 data items were extracted to answer our research question and give us a broad understanding of this field. Extracted data items included study type, sample size, ML/DL task, use of external validation, and ability to explain, as detailed in Table 1. Initially, a set of 50 studies were selected, and their characteristics were extracted by the authors and discussed as a group to level their understanding of the required fields. After reaching a consensus on this subset, the authors independently extracted data from the remaining articles. During this process, any ambiguities were checked with others in a private online forum. Due to a large number of remaining studies, we were not able to assess the bias of the individual articles. Data analysis and plotting were done using Python (v3.9; Python Software Foundation, USA) and the Plotly library (v5.10.0; Plotly Inc., USA) [26].

Table 1. Extracted characteristics.

Characteristic	Description
Year	The year it was published based on MedLine database
Number of Subjects	Binned into categories of <100, 100–1000, 1000–10,000, 10,000–100,000, and >100,000
Country of the Authors	The affiliation country of the corresponding author
Clinical Application Type	Either diagnosis, treatment, prognosis, informatics, combined, or other
Study Modality	Modality or modalities that were used in the study
Studied Organ	The organ that was studied
Studied Disease	If there was a specific pathology that the study targeted
ML Methodology Category	Either conventional machine learning, deep learning, or a combination of both
ML Task Type	Classification, regression, segmentation, object detection, image generation, or other (multi-option)
Use of External Validation	If they used external data to validate their pipeline
Use of Explainable Methods	If they used any explainable method (only for deep-learning-based studies with classification or survival analysis tasks)

In the synthesis section, we did a thematic analysis of the extracted information. For the presented results, we first give some general information on that theme and why it is important and then report our findings.

3. Results and Synthesis

The systematic search yielded 2237 manuscripts. After careful examination, 652 were excluded yielding a final number of 1584 studies that met our eligibility criteria, detailed in Figure 1. From 2012 to 2021, the number of published studies increased year by year, with large increases in 2020 and 2021, as seen in Figure 2. This can be attributed to the COVID-19 pandemic and the resulting influx of publications on this topic [27]. Moreover, based on the first five months, the forecasted number of publications in 2022 was around 780, which is less than the yearly doubling trend seen since 2017, confirming the decrease of the “COVID publication fever”. As shown in Figure 3, China, the United States, and South Korea have the highest number of publications in this field, with 444, 290, and 82 manuscripts, respectively.

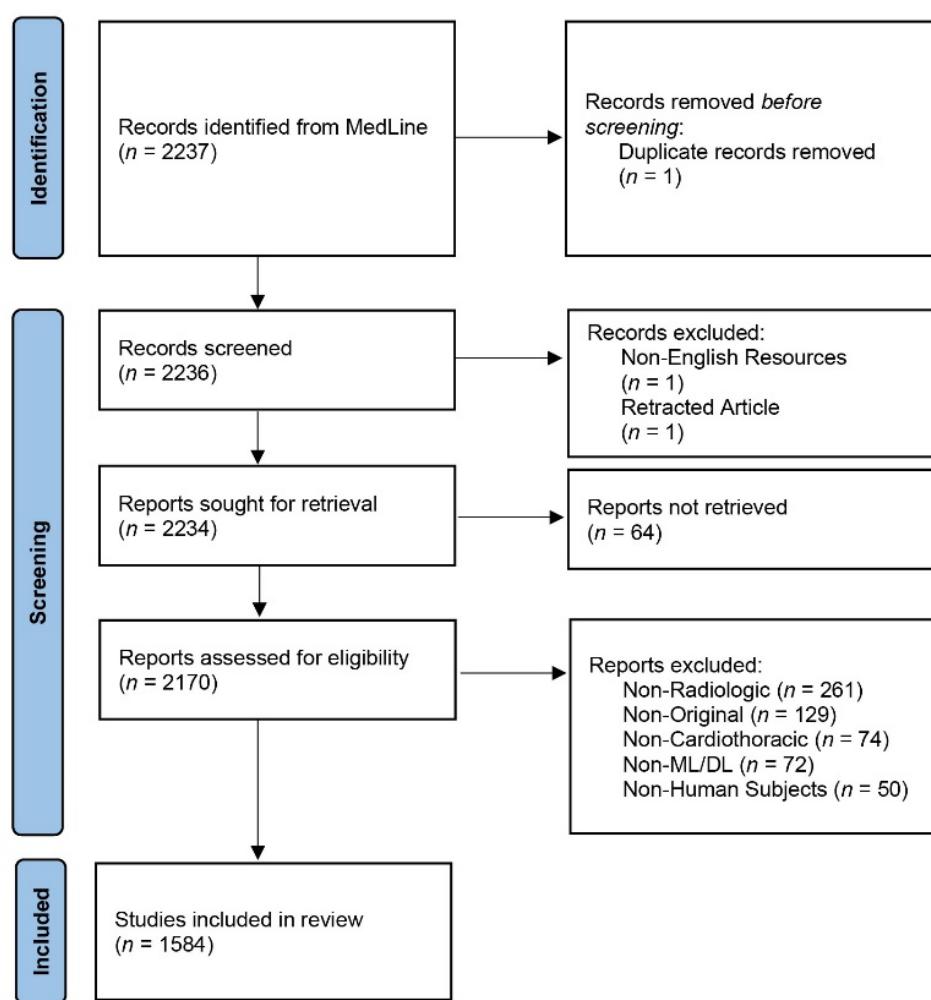


Figure 1. PRISMA inclusion flow chart.

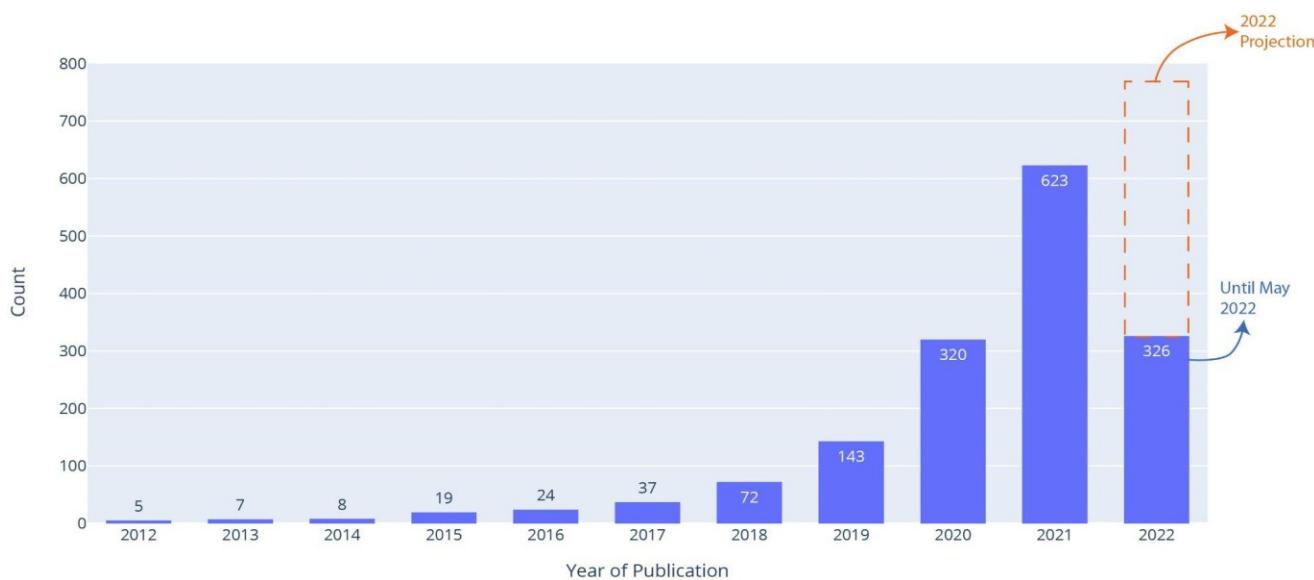


Figure 2. Yearly distribution of included studies.

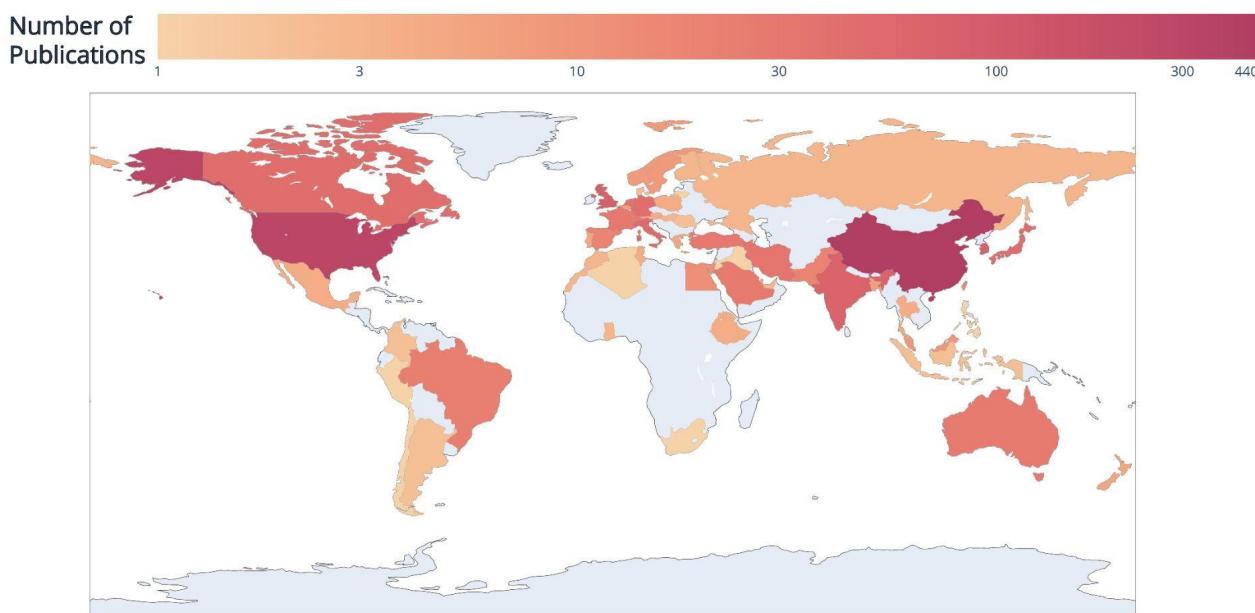


Figure 3. Geographical distribution of included studies. The scale is log-based to better visualize the variability between different countries.

3.1. Clinical Application

Among the included studies, 1000 focused on disease diagnosis (63%) and 123 (8%) studies targeted patient prognosis, as seen in Figure 4. Three hundred sixty-seven studies (23%) worked on organ segmentation or image quality improvement and were categorized as informatics-related. This segmentation can be further utilized in downstream tasks; for example, Qi et al. used a deep-learning algorithm to segment lung nodules and used the masks for longitudinal surveillance of lung-cancer patients [28]. It is noteworthy to mention that if an informatics task, like segmentation, was utilized for a clinical purpose, the study was categorized based on the clinical application.

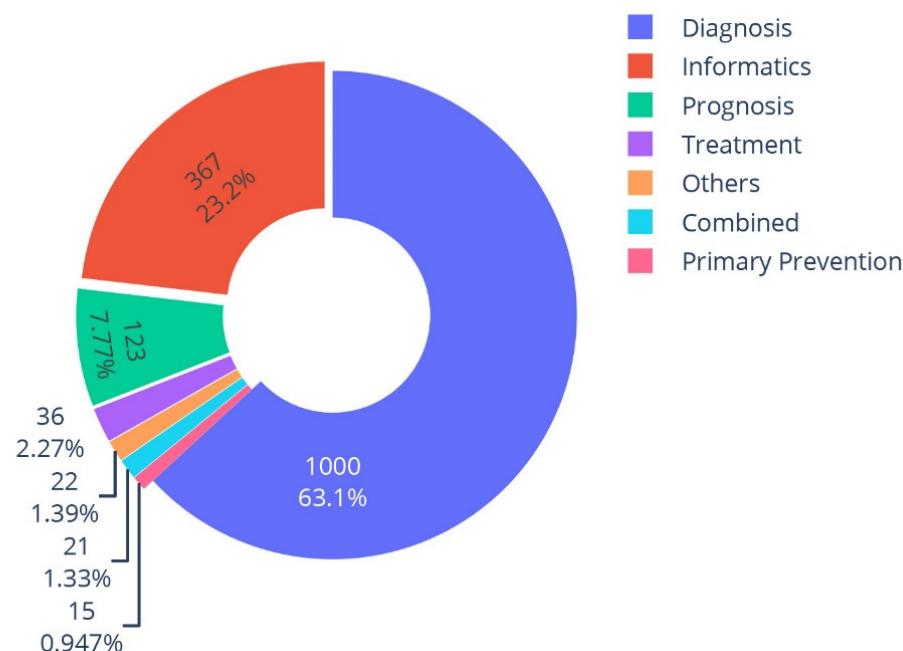


Figure 4. Distribution of clinical applications of the included studies.

3.2. Organ and Pathology

Most of the studies focused on the lung as their primary organ of interest, comprising 1025/1584 (65%), and 514/1584 (32%) of studies worked on the cardiovascular system. Ten studies worked on multiple-organ systems. The majority of studies worked on COVID-19 and lung-nodule detection and classification, as seen in Table 2.

Table 2. Distribution of organs of interest and investigated pathologies. Note that several studies worked on multiple pathologies from different organs.

Organ	Pathology	Count
Lung	COVID-19	551
	Malignancy	265
	Interstitial Lung Diseases	93
	Infection (non-COVID-19)	89
	Obstructive Lung Diseases	82
	Pneumothorax	61
	Pulmonary Edema	58
	Pleural Effusion	56
	Atelectasis	53
	Tuberculosis	10
	Respiratory Distress Syndrome	6
	Cystic Fibrosis	5
	No Specific Pathology	87
Heart	Coronary Artery Disease	114
	Cardiomegaly	90
	Valvular Disorders	28
	Heart Failure	26
	Cardiomyopathy and Myocardial Disease	21
	Arrhythmia	15
	Congenital Heart Diseases	6
	Fat Analysis	5
	Pericarditis	1
	No Specific Pathology	195
Vascular System	Aortic Aneurysm and Dissection	6
	Pulmonary Hypertension	3
	Coarctation of the Aorta	1
	No Specific Pathology	8
Chest Wall	No Specific Pathology	11
Lymphatic System	Malignancy	4
	No Specific Pathology	1
Thymus	Malignancy	2

COVID-19 was the most commonly studied pathology among the reviewed studies. As the pandemic started, pulmonary CT images and chest radiographs were regarded as first-line screening and diagnostic tools [29]. Although soon reverse transcription polymerase chain reaction (RT-PCR) replaced imaging studies as the gold standard of diagnosis, the AI community was very eager to test their algorithms to see how far they could push the limits of COVID-19 diagnosis based on imaging features [30]. Additionally, many studies used ML/DL to prognosticate patients with COVID-19 and predict severe outcomes, like ICU admission or death [31]. Publicly available datasets and coding challenges fueled this enthusiasm by creating a way to benchmark algorithm performance [32–34].

Lung-nodule detection has gained attention since one of the first large-scale medical datasets that was publicly released was the LUNG Nodule Analysis challenge (LUNA) dataset, which contains 888 lung CT series with the exact location of each nodule [21]. The other publicly available dataset is the Lung Image Database Consortium image collection (LIDC-IDRI), which is comprised of 1018 lung CT examinations with each nodule being

segmented by cardiothoracic radiologists and a subset (157 patients) labeled as malignant or benign based on pathology reports [35]. These public datasets have paved the way for non-medical researchers to work on these challenging tasks by providing annotated high-quality data, which can be the most important hurdle in a machine-learning project.

Among studies focusing on the heart as their primary organ of interest, the majority were classified in the realm of informatics. For example, Carbaja-Degante et al. focused on ventricle segmentation on cardiac CT and MRI [36]. Additionally, many studies have tried to calculate the calcium scoring of coronary arteries based on cardiac CTs [37,38].

3.3. Imaging Modality

A majority of the studies used CTs for their image analysis, comprising 760/1651 (46%). As 67 studies used two different modalities, the denominator of the fractions in this section is 1651, i.e., the total number of modalities used, rather than 1584 (the number of included studies), as seen in Figure 5. This might be attributed to the fact that there are many publicly available CT imaging datasets, as discussed previously. Some studies feed slices of the 3D volume to a 2D model, essentially increasing their training data, with the cost of losing volumetric information. To preserve spatial information, one can use a 3D model. As one might think, using a full 3D volume to train a DL model needs a high-performing computational infrastructure. One way to overcome the computational challenge is an approach called *2.5D training*. As an example, Kim et al. first segmented lung nodules in CT images [39]. Then, using 3D regions of interest for the segmented nodules, they obtained nine different 2D images in the axial, coronal, and sagittal planes with two additional oblique cross sections (45° and -45°) per plane. Finally, they combined them into one image that is not truly 3D or 2D; hence, it is called 2.5D. They showed that for the specific problem of differentiating adenocarcinomas, the 2.5D approach outperformed the 3D approach and was comparable to radiologists' performance with much lower computational costs compared to 3D.

X-ray radiographs are the second most common modality used in the included ML/DL studies. This can be attributed to the fact that radiographs are very common in both developed and developing countries, and many diseases can be diagnosed based on this 2D image. Additionally, there are several large-scale public datasets on chest X-rays, namely the CheXpert dataset and the NIH chest X-ray dataset [40,41]. When combined, these two datasets provide more than 300,000 images, each labeled for the presence of several conditions, including pleural effusion, pneumonia, pneumothorax, and nodules.

3.4. Machine-Learning-Algorithm Type and Sample Size

A total of 1217/1584 (77%) studies used deep learning, and only 242/1584 studies (15%) used conventional machine learning for cardiothoracic image analysis; the other 125 (8%) used a combination of conventional machine learning and deep learning, as seen in Figure 6. As expected, studies utilizing DL had a higher number of imaging samples, as the final performance of these models highly depends on the quantity and quality of the training data. This should be noted because many studies only mentioned the number of included patients, which might not always be equal to the number of examinations used for training and might be misleading to their readership.

Conventional machine-learning algorithms are known to work well with tabular data. One popular approach to use these models for image analysis is by utilizing radiomics features [42]. Radiomics is a method for quantitatively describing medical images; in contrast to radiologists who might report a pulmonary nodule as "... a 3 mm perifissural nodular opacity within the lingula ...", radiomics describes the nodule with numerical values contributing to the overall shape, texture, sphericity, contrast with surrounding tissue, etc. [43]. To perform radiomics-feature extraction, one has to segment the region of interest, either manually or by deep-learning algorithms, and run predefined algorithms on that region [44]. The output is a high-dimensional table filled with numerical values, which is the input to a conventional ML model.

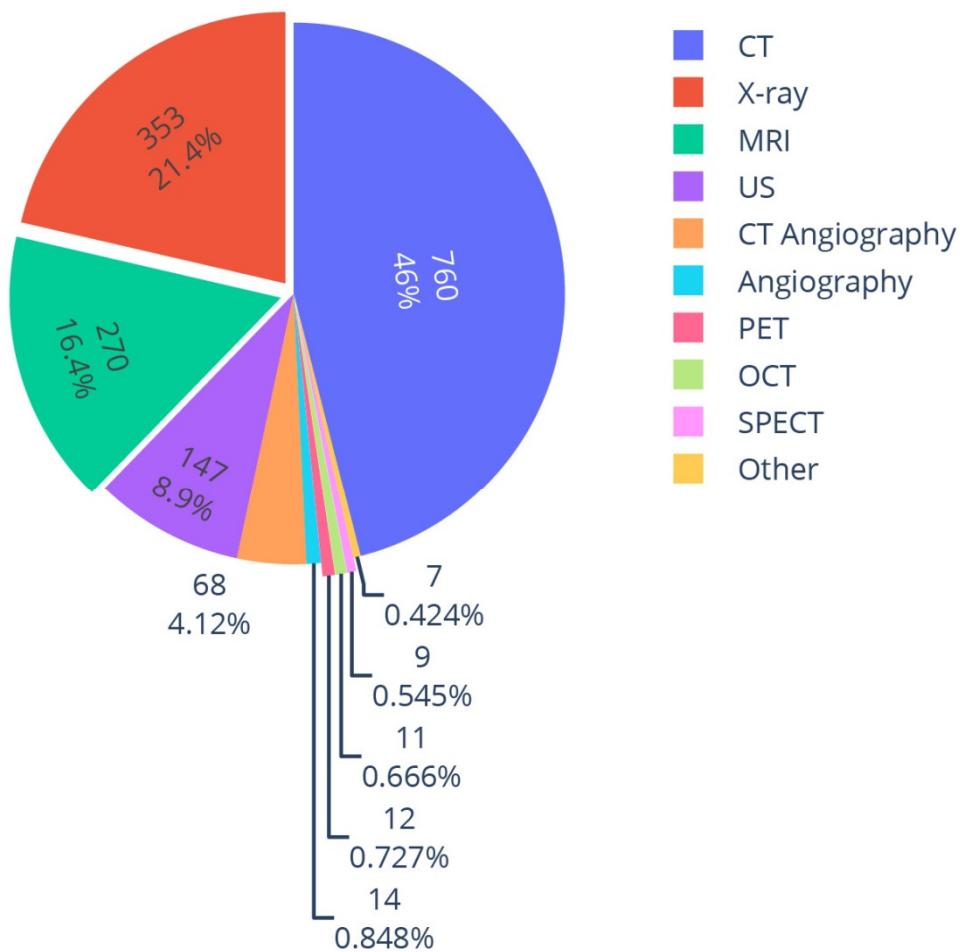


Figure 5. Distribution of target modalities of the included studies.

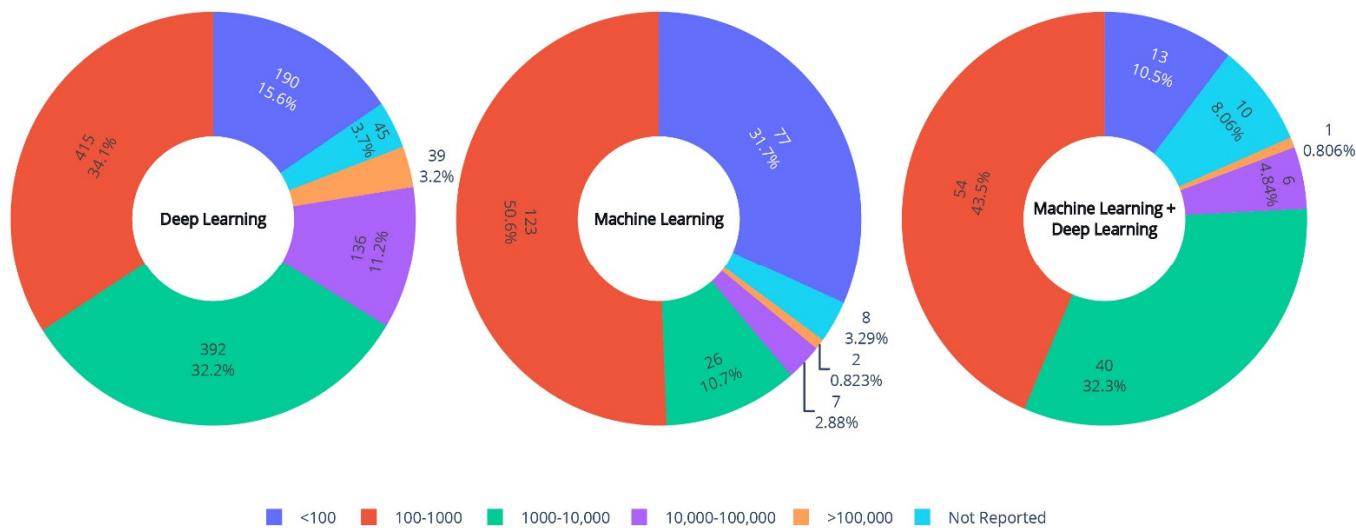


Figure 6. Distribution of different machine-learning methodologies and sample size of the included studies.

One approach that enables using DL algorithms to perform well in limited data settings is a technique called *transfer learning*, in which the parameters of an already-trained DL model are used in a downstream task [45]. There are plenty of publicly available models that are pretrained on natural or medical images that can be used for feature extraction for

downstream tasks [46,47]. Similarly, Chen et al. used such a pretrained DL model, applied the model to endobronchial ultrasound images (without retraining), extracted features, and used a machine-learning algorithm to predict malignancy of the pulmonary nodules [48]. This technique enabled them to reach an area under the receiving operating curve (AUROC) of 0.87 on only 164 patients.

3.5. Machine Learning Tasks

ML and DL tasks can be divided into five major categories.

3.5.1. Classification

The goal of classification is to assign a single label to the whole image, for example, whether a chest X-ray shows signs of pneumonia or not. These tasks can be binary (i.e., yes/no labels), multiclass (selecting one option from more than two choices that are mutually exclusive, e.g., viral pneumonia/bacterial pneumonia/COVID), or multi-label (where there is potential for the co-occurrence of conditions, e.g., having both cardiomegaly and pleural effusion) [49–52]. Of the included studies, 869 used a classification model in a part of their pipelines. Of these, 190 studies combined classification with other types of ML tasks, for example, first segmenting a lung nodule and then classifying the isolated nodule as benign or malignant [53].

3.5.2. Regression

Regression tasks generate a continuous numerical value from the input data. A common example is predicting the survival of patients with lung cancer; it should be noted that by using noncontinuous time periods in survival analysis, the task becomes a multiclass classification [54]. Eighty-eight studies used regression in their analysis pipelines. Other use cases of regression tasks in cardiothoracic imaging include measurement of aortic calcium score, ventricle volume, and pulmonary lesion volume [55–58].

3.5.3. Object Detection

Sixty-six studies did object detection, which is localizing an object of interest in an image using key points or bounding boxes. For example, Rafael-Palou et al. used a deep-learning network to localize lung nodules and then used this model for automated follow-up of patients with pulmonary nodules [59]. In another example, Pezzano et al. trained an object-detection deep-learning-based model to determine the location of COVID-19 opacities, which can further be used for calculating severity scores for patients [60].

3.5.4. Semantic Segmentation

Semantic segmentation involves exact delineation of the organ of interest by an ML model, and was performed in 529 studies. To highlight the power of deep-learning algorithms in segmentation tasks, Nardelli et al. used a DL model to separate arteries and veins in pulmonary vasculature in CT scans [61]. Other instances of segmentation in cardiothoracic imaging involves delineating ventricular myocardium or valve leaflets that can be used to calculate structural and flow-related parameters [62–65]. Segmentation tasks require experts to annotate the organ of interest, which makes preparing the training data both time-consuming and expensive. Some techniques can help with reducing the required training data without hurting the model performance, like Few-Shot Learning. For example, Wang et al. used this technique to achieve a very accurate heart segmentation model with only four annotated CT angiograms [66].

3.5.5. Generative Tasks

In this broad category of ML tasks, the model is used to create images similar to, but not exactly the same as, the original input data. For example, these models have been used to increase image resolution or create a dual-energy X-ray from traditional X-rays [67,68]. Overall, 125 studies used generative approaches in their methodology design. An intuitive

use case of generative models is artificially increasing low prevalence classes in the dataset by creating synthetic images with those characteristics. Astaraki et al. used this technique to create synthetic images of pulmonary nodules based on a limited dataset in order to train a segmentation model [69]. They showed that the trained segmentation model performs well on real patient images, highlighting the power of DL-based image generation. Another popular application of generative models in cardiothoracic imaging is creating standard-dose scans from low-dose sequences [70,71].

3.6. External Validation

The most important issue that hinders ML/DL algorithm adoption in healthcare is their lack of generalizability, meaning that an algorithm can perform well on a particular set of data (similar to the one it was trained on) while having suboptimal performance on external data [72]. This potential drawback necessitates rigorous validation of these algorithms on external data, called external validation. Without this process, one cannot trust if the reported performance of the model can happen in other real clinical cases [73,74]. Of the reviewed studies, 245 (15%) utilized external validation to test the generalizability of their models. Source that can be used for external validation are the publicly available datasets; in which case, the algorithm is trained on the institutional data, and then its performance is evaluated on the public set, allowing fair comparison between the different algorithms [75–77].

3.7. Interpretability Maps

DL algorithms are prone to biases that are caused by the model taking “shortcuts” rather than focusing on important features. In an interesting study, Rueckel et al. evaluated this phenomenon for two publicly available pneumothorax-detection algorithms [78]. They have shown that these algorithms actually detect the inserted chest tube on the X-rays as a surrogate marker for pneumothorax, causing substantial performance differences in patients with and without chest tubes. This bias can especially be present in many DL-based classification/regression algorithms, as they give a label to the whole image, but the basis for its decisions is often difficult to obtain. Therefore, these DL models are sometimes referred to as black boxes.

There are some ways that researchers can gain insight into how DL networks reach a particular decision, which is called interpretability maps [79]. Though these maps can help with the adoption of DL models in clinical settings, they are not perfect themselves and might cause biases [80]. Teng et al. provide a decent overview of how these tools work and how they can be used in medical image analysis [81]. Further, they only provide location information and not the reason that location was selected. Of the 694 studies that used only classification or regression in their image-analysis pipelines, 184 presented some visual explanations of how the model reached its decision.

4. Discussion

Artificial intelligence, and specifically deep learning, is a rapidly evolving field of computer science. As these algorithms can help with task automation, they have also gained attraction in the medical domain. In this scoping review, we aimed to provide a holistic view of ML and DL applications in cardiothoracic imaging, while summarizing some important nuances of this field. As evidenced by our findings, there is a growing number of publications in cardiothoracic imaging, and the COVID-19 pandemic sparked a big jump in the total number of studies. Researchers have used ML/DL to solve many clinical problems, with a higher focus on disease diagnosis. Having said this, some factors hinder the adoption of these algorithms in clinical practice. Based on this review, here are some future directions that ML researchers can pursue in the field of cardiothoracic imaging.

4.1. Generalizability Testing

Although a limited number of the reviewed studies externally validated their algorithm, most of the included studies did not test the generalizability of their models on external data. This is a crucial step toward showing a proof-of-concept study that can be useful in real clinical scenarios. It is of utmost importance to use data from different sources as the training data, not just setting aside a portion of training data for testing purposes, like many of the reviewed studies did. This causes inherent biases in the training data being present in the test set, causing a misleadingly high performance [82]. One solution would be using publicly available datasets as a community-assigned benchmark. Although this means refraining from using these datasets during training, it ensures comparable results across different studies, but it would also demand these public datasets be representative of the populations where the algorithm is deployed.

4.2. Applied Research

As many radiologists believe, AI is here to help them with their day-to-day clinical workflows, rather than replacing them. However, not many studies tested these algorithms side-by-side to a radiologist, and those who did compared them against each other, rather than in a cooperative-reading setting [83,84]. For AI to be integrated into the clinical workflow, extensive studies are needed to quantify the AI's help and to show its efficiency and efficacy in prognostic and diagnostic workflows. As an example, it is controversial whether a radiologist should read a study and then see the AI's prediction, or the prediction results should be shown to the radiologist in the first place [85]. These questions should be carefully investigated by researchers in order to facilitate AI adoption.

Our study is subject to several limitations that need to be addressed in future efforts. First, as we wanted to give a general overview of the field, we only used organ keywords from Medical Subject Headings (MeSH terms) in our search query. Such a strategy might have caused us to miss some studies, but we believe our search is still representative of the whole literature. Second, we only searched MedLine as the main database of medical literature; however, many AI researchers might release non-peer-reviewed versions of their work on preprint servers, like arXiv or medRxiv. Finally, because there was a need to include numerous studies to reach the goal of this scoping review, we could not assess the biases of the included papers, as it would make the task intractable. We strongly encourage future studies focusing on a specialized subfield of cardiothoracic imaging to identify hidden biases in ML/DL research in this field.

5. Conclusions

Overall, ML and DL have gained attraction in cardiothoracic imaging research, especially after the COVID pandemic. Researchers have utilized these techniques to analyze a variety of modalities in order to diagnose diseases, monitor treatment, enhance imaging quality, and predict patient prognosis. As a rule of thumb, DL models require more training data compared to other conventional ML algorithms; however, there are some techniques like transfer learning or Few-Shot Learning that can help alleviate this problem. There are also ML designs, like radiomics studies that extract handcrafted features from medical images or creating tabular data ready to be fed to an ML model, which can be done in data-constraint settings. With all these advances, there are some less-explored areas, like testing generalizability, quantifying model uncertainty, and evaluating the effect of AI in radiologists' workflows and decision-making processes that need to be further investigated by ML researchers. While there have been tremendous advances in AI technology in the recent past, it is critical that AI researchers recognize the potential pitfalls of both the AI technology itself and the challenges of integrating these tools into clinical practice.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/diagnostics12102512/s1>, Table S1: Different parts of the search term that were used for retrieving relevant studies.

Author Contributions: Conceptualization, B.K., P.R. and B.J.E.; Methodology, B.K. and P.R.; Software, B.K. and P.R.; Formal Analysis, B.K. and P.R.; Data Curation, B.K., P.R., S.F., M.M., S.V. and E.M.; Writing—Original Draft Preparation, B.K. and P.R.; Writing—Review and Editing, B.K., P.R., S.F., M.M., S.V., E.M., H.C. and B.J.E.; Visualization, B.K.; Supervision, B.J.E.; Project Administration, P.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding, and all authors contributed voluntarily.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data can be shared upon reasonable request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [[CrossRef](#)]
2. Kataria, S.; Ravindran, V. Electronic Health Records: A Critical Appraisal of Strengths and Limitations. *J. R. Coll. Physicians Edinb.* **2020**, *50*, 262–268. [[CrossRef](#)] [[PubMed](#)]
3. Goldstein, B.A.; Navar, A.M.; Carter, R.E. Moving beyond regression techniques in cardiovascular risk prediction: Applying machine learning to address analytic challenges. *Eur. Heart J.* **2016**, *38*, 1805–1814. [[CrossRef](#)]
4. Kersting, K. Machine Learning and Artificial Intelligence: Two Fellow Travelers on the Quest for Intelligent Behavior in Machines. *Front. Big Data* **2018**, *1*, 6. [[CrossRef](#)] [[PubMed](#)]
5. Lee, J.-G.; Jun, S.; Cho, Y.-W.; Lee, H.; Kim, G.B.; Seo, J.B.; Kim, N. Deep Learning in Medical Imaging: General Overview. *Korean J. Radiol.* **2017**, *18*, 570–584. [[CrossRef](#)] [[PubMed](#)]
6. Wan, L.; Zeiler, M.; Zhang, S.; Le Cun, Y.; Fergus, R. Regularization of Neural Networks using DropConnect. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 17–19 June 2013; Dasgupta, S., McAllester, D., Eds.; *Proceedings of Machine Learning Research*. JMLR: Cambridge MA, USA, 2013; Volume 28, pp. 1058–1066. Available online: <https://proceedings.mlr.press/v28/wan13.html> (accessed on 10 September 2022).
7. Taigman, Y.; Yang, M.; Ranzato, M.A.; Wolf, L. Deepface: Closing the Gap to Human-Level Performance in Face Verification. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; Available online: http://openaccess.thecvf.com/content_cvpr_2014/html/Taigman_DeepFace_Closing_the_2014_CVPR_paper.html (accessed on 10 September 2022).
8. Liu, Z.; Hu, H.; Lin, Y.; Yao, Z.; Xie, Z.; Wei, Y.; Ning, J.; Cao, Y.; Zhang, Z.; Dong, L.; et al. Swin Transformer V2: Scaling Up Capacity and Resolution. *arXiv* **2021**, arXiv:2111.09883. Available online: <http://arxiv.org/abs/2111.09883> (accessed on 10 September 2022).
9. Papanicolas, I.; Woskie, L.; Jha, A.K. Health Care Spending in the United States and Other High-Income Countries. *JAMA* **2018**, *319*, 1024–1039. [[CrossRef](#)]
10. Shin, Y.; Kim, S.; Lee, Y.H. AI musculoskeletal clinical applications: How can AI increase my day-to-day efficiency? *Skelet. Radiol.* **2021**, *51*, 293–304. [[CrossRef](#)]
11. Korot, E.; Guan, Z.; Ferraz, D.; Wagner, S.K.; Zhang, G.; Liu, X.; Faes, L.; Pontikos, N.; Finlayson, S.G.; Khalid, H.; et al. Code-free deep learning for multi-modality medical image classification. *Nat. Mach. Intell.* **2021**, *3*, 288–298. [[CrossRef](#)]
12. Santomartino, S.M.; Yi, P.H. Systematic Review of Radiologist and Medical Student Attitudes on the Role and Impact of AI in Radiology. *Acad. Radiol.* **2022**, *in press*. [[CrossRef](#)]
13. Huisman, M.; Ranschaert, E.; Parker, W.; Mastrodicasa, D.; Koci, M.; de Santos, D.P.; Coppola, F.; Morozov, S.; Zins, M.; Bohyn, C.; et al. An international survey on AI in radiology in 1,041 radiologists and radiology residents part 1: Fear of replacement, knowledge, and attitude. *Eur. Radiol.* **2021**, *31*, 7058–7066. [[CrossRef](#)] [[PubMed](#)]
14. Ooi, S.; Makmur, A.; Soon, Y.; Fook-Chong, S.; Liew, C.; Sia, D.; Ting, Y.; Lim, C. Attitudes toward artificial intelligence in radiology with learner needs assessment within radiology residency programmes: A national multi-programme survey. *Singap. Med. J.* **2021**, *62*, 126–134. [[CrossRef](#)]
15. Collado-Mesa, F.; Alvarez, E.; Arheart, K. The Role of Artificial Intelligence in Diagnostic Radiology: A Survey at a Single Radiology Residency Training Program. *J. Am. Coll. Radiol.* **2018**, *15*, 1753–1757. [[CrossRef](#)]
16. Eltorai, A.E.; Bratt, A.K.; Guo, H.H. Thoracic Radiologists’ Versus Computer Scientists’ Perspectives on the Future of Artificial Intelligence in Radiology. *J. Thorac. Imaging* **2019**, *35*, 255–259. [[CrossRef](#)] [[PubMed](#)]
17. Scheetz, J.; Rothschild, P.; McGuinness, M.; Hadoux, X.; Soyer, H.P.; Janda, M.; Condon, J.J.; Oakden-Rayner, L.; Palmer, L.J.; Keel, S.; et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. *Sci. Rep.* **2021**, *11*, 5193. [[CrossRef](#)]

18. Gong, B.; Nugent, J.P.; Guest, W.; Parker, W.; Chang, P.J.; Khosa, F.; Nicolaou, S. Influence of Artificial Intelligence on Canadian Medical Students’ Preference for Radiology Specialty: ANational Survey Study. *Acad. Radiol.* **2018**, *26*, 566–577. [CrossRef] [PubMed]
19. Bin Dahmash, A.; Alabdulkareem, M.; Alfutais, A.; Kamel, A.M.; Alkholaiwi, F.; AlShehri, S.; Al Zahrani, Y.; Almoaiqel, M. Artificial intelligence in radiology: Does it impact medical students preference for radiology as their future career? *BJR | Open* **2020**, *2*, 20200037. [CrossRef] [PubMed]
20. Sorantin, E.; Grassler, M.G.; Hemmelmayr, A.; Tschauner, S.; Hrzic, F.; Weiss, V.; Lacekova, J.; Holzinger, A. The augmented radiologist: Artificial intelligence in the practice of radiology. *Pediatr. Radiol.* **2021**, *52*, 2074–2086. [CrossRef]
21. Setio, A.A.A.; Traverso, A.; de Bel, T.; Berens, M.S.; Bogaard, C.V.D.; Cerello, P.; Chen, H.; Dou, Q.; Fantacci, M.E.; Geurts, B.; et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge. *Med. Image Anal.* **2017**, *42*, 1–13. [CrossRef]
22. Abiodun, K.M.; Awotunde, J.B.; Aremu, D.R.; Adeniyi, E.A. Explainable AI for Fighting COVID-19 Pandemic: Opportunities, Challenges, and Future Prospects. In *Computational Intelligence for COVID-19 and Future Pandemics*; Springer: Singapore, 2022; pp. 315–332. [CrossRef]
23. Tricco, A.C.; Lillie, E.; Zarin, W.; O’Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.J.; Horsley, T.; Weeks, L.; et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [CrossRef]
24. Peng, Y.; Liu, E.; Peng, S.; Chen, Q.; Li, D.; Lian, D. Using artificial intelligence technology to fight COVID-19: A review. *Artif. Intell. Rev.* **2022**, *55*, 4941–4977. [CrossRef] [PubMed]
25. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *NIPS* **2017**, *60*, 84–90. [CrossRef]
26. Van Rossum, G.; Drake, F.L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, USA, 2009; Available online: <https://dl.acm.org/doi/abs/10.5555/1593511> (accessed on 10 September 2022).
27. Raynaud, M.; Goutaudier, V.; Louis, K.; Al-Awadhi, S.; Dubourg, Q.; Truchot, A.; Brousse, R.; Saleh, N.; Giarraputo, A.; Debiais, C.; et al. Impact of the COVID-19 pandemic on publication dynamics and non-COVID-19 research production. *BMC Med. Res. Methodol.* **2021**, *21*, 255. [CrossRef] [PubMed]
28. Qi, L.-L.; Wang, J.-W.; Yang, L.; Huang, Y.; Zhao, S.-J.; Tang, W.; Jin, Y.-J.; Zhang, Z.-W.; Zhou, Z.; Yu, Y.-Z.; et al. Natural history of pathologically confirmed pulmonary subsolid nodules with deep learning-assisted nodule segmentation. *Eur. Radiol.* **2020**, *31*, 3884–3897. [CrossRef] [PubMed]
29. Fang, Y.; Zhang, H.; Xie, J.; Lin, M.; Ying, L.; Pang, P.; Ji, W. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* **2020**, *296*, 2020200432. [CrossRef] [PubMed]
30. Summers, R.M. Artificial Intelligence of COVID-19 Imaging: A Hammer in Search of a Nail. *Radiology* **2021**, *298*, E162–E164. [CrossRef]
31. Shiri, I.; Sorouri, M.; Geramifar, P.; Nazari, M.; Abdollahi, M.; Salimi, Y.; Khosravi, B.; Askari, D.; Aghaghazvini, L.; Hajianfar, G.; et al. Machine learning-based prognostic modeling using clinical data and quantitative radiomic features from chest CT images in COVID-19 patients. *Comput. Biol. Med.* **2021**, *132*, 104304. [CrossRef]
32. Shih, G.; Wu, C.C.; Halabi, S.S.; Kohli, M.D.; Prevedello, L.M.; Cook, T.S.; Sharma, A.; Amorosa, J.K.; Arteaga, V.; Galperin-Aizenberg, M.; et al. Augmenting the National Institutes of Health Chest Radiograph Dataset with Expert Annotations of Possible Pneumonia. *Radiol. Artif. Intell.* **2019**, *1*, e180041. [CrossRef]
33. de la Iglesia Vayá, M.; Saborit, J.M.; Montell, J.A.; Pertusa, A.; Bustos, A.; Cazorla, M.; Galant, J.; Barber, X.; Orozco-Beltrán, D.; García-García, F.; et al. BIMCV COVID-19+: A large annotated dataset of RX and CT images from COVID-19 patients. *arXiv* **2020**, arXiv:2006.01174. Available online: <http://arxiv.org/abs/2006.01174> (accessed on 10 September 2022).
34. Lakhani, P.; Mongan, J.; Singhal, C.; Zhou, Q.; Andriole, K.P.; Auffermann, W.F.; Prasanna, P.; Pham, T.; Peterson, M.; Bergquist, P.J.; et al. The 2021 SIIM-FISABIO-RSNA Machine Learning COVID-19 Challenge: Annotation and Standard Exam Classification of COVID-19 Chest Radiographs. *2021*. Available online: <https://osf.io/532ek> (accessed on 10 September 2022).
35. Armato, S.G.; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.R.; Henschke, C.I.; Hoffman, E.; et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. *Med. Phys.* **2011**, *38*, 915–931. [CrossRef]
36. Carballo-Degante, E.; Avendaño, S.; Ledesma, L.; Olveres, J.; Vallejo, E.; Escalante-Ramirez, B. A multiphase texture-based model of active contours assisted by a convolutional neural network for automatic CT and MRI heart ventricle segmentation. *Comput. Methods Programs Biomed.* **2021**, *211*, 106373. [CrossRef] [PubMed]
37. Lee, J.-G.; Kim, H.; Kang, H.; Koo, H.J.; Kang, J.-W.; Kim, Y.-H.; Yang, D.H. Fully Automatic Coronary Calcium Score Software Empowered by Artificial Intelligence Technology: Validation Study Using Three CT Cohorts. *Korean J. Radiol.* **2021**, *22*, 1764. [CrossRef]
38. Lee, S.; Rim, B.; Jou, S.-S.; Gil, H.-W.; Jia, X.; Lee, A.; Hong, M. Deep-Learning-Based Coronary Artery Calcium Detection from CT Image. *Sensors* **2021**, *21*, 7059. [CrossRef]
39. Kim, H.; Lee, D.; Cho, W.S.; Lee, J.C.; Goo, J.M.; Kim, H.C.; Park, C.M. CT-based deep learning model to differentiate invasive pulmonary adenocarcinomas appearing as subsolid nodules among surgical candidates: Comparison of the diagnostic performance with a size-based logistic model and radiologists. *Eur. Radiol.* **2020**, *30*, 3295–3305. [CrossRef] [PubMed]

40. Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R.M. ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2097–2106.
41. Irvin, J.; Rajpurkar, P.; Ko, M.; Yu, Y.; Ciurea-Ilcus, S.; Chute, C.; Marklund, H.; Haghgoo, B.; Ball, R.; Shpanskaya, K.; et al. CheXpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *arXiv* **2019**, arXiv:1901.07031. Available online: <http://arxiv.org/abs/1901.07031> (accessed on 10 September 2022). [CrossRef]
42. Grinsztajn, L.; Oyallon, E.; Varoquaux, G. Why do tree-based models still outperform deep learning on tabular data? *arXiv* **2022**, arXiv:2207.08815. Available online: <http://arxiv.org/abs/2207.08815> (accessed on 10 September 2022).
43. Rizzo, S.; Botta, F.; Raimondi, S.; Origgi, D.; Fanciullo, C.; Morganti, A.G.; Bellomi, M. Radiomics: The facts and the challenges of image analysis. *Eur. Radiol. Exp.* **2018**, *2*, 36. [CrossRef]
44. Van Timmeren, J.E.; Cester, D.; Tanadini-Lang, S.; Alkadhi, H.; Baessler, B. Radiomics in medical imaging—“how-to” guide and critical reflection. *Insights Imaging* **2020**, *11*, 91. [CrossRef]
45. Zhang, K.; Khosravi, B.; Vahdati, S.; Faghani, S.; Nugen, F.; Rassoulinejad-Mousavi, S.M.; Moassefi, M.; Jagtap, J.M.M.; Singh, Y.; Rouzrokh, P.; et al. Mitigating Bias in Radiology Machine Learning: 2. Model Development. *Radiol. Artif. Intell.* **2022**, *4*, e220010. [CrossRef]
46. Wightman, R.; Soare, A.; Arora, A.; Ha, C.; Raw, N.; Mike; Chen, R.; Rizin, M.; Kim, H.; Kertész, C.; et al. Wrightman/Pytorch-Image-Models: TPU VM Trained Weight Release w/PyTorch XLA. 2022. Available online: <https://zenodo.org/record/6369353> (accessed on 10 September 2022).
47. The MONAI Consortium. Project MONAI. 2020. Available online: <https://zenodo.org/record/4323059> (accessed on 10 September 2022).
48. Chen, C.-H.; Lee, Y.-W.; Huang, Y.-S.; Lan, W.-R.; Chang, R.-F.; Tu, C.-Y.; Chen, C.-Y.; Liao, W.-C. Computer-aided diagnosis of endobronchial ultrasound images using convolutional neural network. *Comput. Methods Programs Biomed.* **2019**, *177*, 175–182. [CrossRef]
49. Ortiz-Toro, C.; García-Pedrero, A.; Lillo-Saavedra, M.; Gonzalo-Martín, C. Automatic detection of pneumonia in chest X-ray images using textural features. *Comput. Biol. Med.* **2022**, *145*, 105466. [CrossRef]
50. Clark, A.R.; Her, E.J.; Metcalfe, R.; Byrnes, C.A. Could automated analysis of chest X-rays detect early bronchiectasis in children? *Eur. J. Pediatr.* **2021**, *180*, 3171–3179. [CrossRef] [PubMed]
51. Zhu, L.; Xu, Z.; Fang, T. Analysis of Cardiac Ultrasound Images of Critically Ill Patients Using Deep Learning. *J. Healthc. Eng.* **2021**, *2021*, 6050433. [CrossRef]
52. Cahan, N.; Marom, E.M.; Soffer, S.; Barash, Y.; Konen, E.; Klang, E.; Greenspan, H. Weakly supervised attention model for RV strain classification from volumetric CTPA scans. *Comput. Methods Programs Biomed.* **2022**, *220*, 106815. [CrossRef]
53. Kasinathan, G.; Jayakumar, S. Cloud-Based Lung Tumor Detection and Stage Classification Using Deep Learning Techniques. *BioMed Res. Int.* **2022**, *2022*, 4185835. [CrossRef]
54. Ninomiya, K.; Arimura, H. Homological radiomics analysis for prognostic prediction in lung cancer patients. *Phys. Medica* **2020**, *69*, 90–100. [CrossRef]
55. Pu, J.; Sechrist, J.; Meng, X.; Leader, J.K.; Sciurba, F.C. A pilot study: Quantify lung volume and emphysema extent directly from two-dimensional scout images. *Med. Phys.* **2021**, *48*, 4316–4325. [CrossRef] [PubMed]
56. Guilenea, F.N.; Casciaro, M.E.; Pascaner, A.F.; Soulat, G.; Mousseaux, E.; Craiem, D. Thoracic Aorta Calcium Detection and Quantification Using Convolutional Neural Networks in a Large Cohort of Intermediate-Risk Patients. *Tomography* **2021**, *7*, 636–649. [CrossRef] [PubMed]
57. Winkelmann, M.T.; Jacoby, J.; Schwemmer, C.; Faby, S.; Krumm, P.; Artzner, C.; Bongers, M.N. Fully Automated Artery-Specific Calcium Scoring Based on Machine Learning in Low-Dose Computed Tomography Screening. *Rofo* **2022**, *194*, 763–770. [CrossRef]
58. Chen, R.; Xu, C.; Dong, Z.; Liu, Y.; Du, X. DeepCQ: Deep multi-task conditional quantification network for estimation of left ventricle parameters. *Comput. Methods Programs Biomed.* **2020**, *184*, 105288. [CrossRef]
59. Rafael-Palou, X.; Aubanell, A.; Bonavita, I.; Ceresa, M.; Piella, G.; Ribas, V.; Ballester, M.G. Re-Identification and growth detection of pulmonary nodules without image registration using 3D siamese neural networks. *Med. Image Anal.* **2020**, *67*, 101823. [CrossRef]
60. Pezzano, G.; Díaz, O.; Ripoll, V.R.; Radeva, P. CoLe-CNN+: Context learning—Convolutional neural network for COVID-19-Ground-Glass-Opacities detection and segmentation. *Comput. Biol. Med.* **2021**, *136*, 104689. [CrossRef] [PubMed]
61. Nardelli, P.; Jimenez-Carretero, D.; Bermejo-Pelaez, D.; Washko, G.R.; Rahaghi, F.N.; Ledesma-Carbayo, M.J.; Estepar, R.S.J. Pulmonary Artery–Vein Classification in CT Images Using Deep Learning. *IEEE Trans. Med. Imaging* **2018**, *37*, 2428–2440. [CrossRef]
62. Corinzia, L.; Laumer, F.; Candreva, A.; Taramasso, M.; Maisano, F.; Buhmann, J.M. Neural collaborative filtering for unsupervised mitral valve segmentation in echocardiography. *Artif. Intell. Med.* **2020**, *110*, 101975. [CrossRef]
63. Astudillo, P.; De Beule, M.; Dambre, J.; Mortier, P. Towards safe and efficient preoperative planning of transcatheter mitral valve interventions. *Morphologie* **2019**, *103*, 139–147. [CrossRef]
64. Guo, B.J.; He, X.; Lei, Y.; Harms, J.; Wang, T.; Curran, W.J.; Liu, T.; Zhang, L.J.; Yang, X. Automated left ventricular myocardium segmentation using 3D deeply supervised attention U-net for coronary computed tomography angiography; CT myocardium segmentation. *Med. Phys.* **2020**, *47*, 1775–1785. [CrossRef]

65. Wang, K.-N.; Yang, X.; Miao, J.; Li, L.; Yao, J.; Zhou, P.; Xue, W.; Zhou, G.-Q.; Zhuang, X.; Ni, D. AWSnet: An auto-weighted supervision attention network for myocardial scar and edema segmentation in multi-sequence cardiac magnetic resonance images. *Med. Image Anal.* **2022**, *77*, 102362. [[CrossRef](#)]
66. Wang, W.; Xia, Q.; Hu, Z.; Yan, Z.; Li, Z.; Wu, Y.; Huang, N.; Gao, Y.; Metaxas, D.; Zhang, S. Few-Shot Learning by a Cascaded Framework With Shape-Constrained Pseudo Label Assessment for Whole Heart Segmentation. *IEEE Trans. Med. Imaging* **2021**, *40*, 2629–2641. [[CrossRef](#)]
67. Lee, D.; Kim, H.; Choi, B.; Kim, H.-J. Development of a deep neural network for generating synthetic dual-energy chest x-ray images with single x-ray exposure. *Phys. Med. Biol.* **2019**, *64*, 115017. [[CrossRef](#)] [[PubMed](#)]
68. Gomi, T.; Hara, H.; Watanabe, Y.; Mizukami, S. Improved digital chest tomosynthesis image quality by use of a projection-based dual-energy virtual monochromatic convolutional neural network with super resolution. *PLoS ONE* **2020**, *15*, e0244745. [[CrossRef](#)]
69. Astaraki, M.; Smedby, Ö.; Wang, C. Prior-aware autoencoders for lung pathology segmentation. *Med. Image Anal.* **2022**, *80*, 102491. [[CrossRef](#)] [[PubMed](#)]
70. Liu, J.; Zhang, Y.; Zhao, Q.; Lv, T.; Wu, W.; Cai, N.; Quan, G.; Yang, W.; Chen, Y.; Luo, L.; et al. Deep iterative reconstruction estimation (DIRE): Approximate iterative reconstruction estimation for low dose CT imaging. *Phys. Med. Biol.* **2019**, *64*, 135007. [[CrossRef](#)] [[PubMed](#)]
71. Wu, J.; Dai, F.; Hu, G.; Mou, X. Low dose CT reconstruction via L1 norm dictionary learning using alternating minimization algorithm and balancing principle. *J. X-ray Sci. Technol.* **2018**, *26*, 603–622. [[CrossRef](#)] [[PubMed](#)]
72. Faghani, S.; Khosravi, B.; Zhang, K.; Moassefi, M.; Jagtap, J.M.; Nugen, F.; Vahdati, S.; Kuanar, S.P.; Rassoulinejad-Mousavi, S.M.; Singh, Y.; et al. Mitigating Bias in Radiology Machine Learning: 3. Performance Metrics. *Radiol. Artif. Intell.* **2022**, *4*, e220061. [[CrossRef](#)]
73. Park, S.H. Diagnostic Case-Control versus Diagnostic Cohort Studies for Clinical Validation of Artificial Intelligence Algorithm Performance. *Radiology* **2019**, *290*, 272–273. [[CrossRef](#)]
74. Yu, A.C.; Eng, J. One Algorithm May Not Fit All: How Selection Bias Affects Machine Learning Performance. *RadioGraphics* **2020**, *40*, 1932–1937. [[CrossRef](#)] [[PubMed](#)]
75. Garau, N.; Paganelli, C.; Summers, P.; Bassis, D.; Lanza, C.; Minotti, M.; De Fiori, E.; Baroni, G.; Rampinelli, C. A segmentation tool for pulmonary nodules in lung cancer screening: Testing and clinical usage. *Phys. Medica* **2021**, *90*, 23–29. [[CrossRef](#)]
76. Heuvelmans, M.A.; van Ooijen, P.M.; Ather, S.; Silva, C.F.; Han, D.; Heussel, C.P.; Hickes, W.; Kauczor, H.-U.; Novotny, P.; Peschl, H.; et al. Lung cancer prediction by Deep Learning to identify benign lung nodules. *Lung Cancer* **2021**, *154*, 1–4. [[CrossRef](#)]
77. Dong, S.; Pan, Z.; Fu, Y.; Yang, Q.; Gao, Y.; Yu, T.; Shi, Y.; Zhuo, C. DeU-Net 2.0: Enhanced deformable U-Net for 3D cardiac cine MRI segmentation. *Med. Image Anal.* **2022**, *78*, 102389. [[CrossRef](#)]
78. Rueckel, J.; Trappmann, L.; Schachtner, B.; Wesp, P.; Hoppe, B.F.; Fink, N.; Ricke, J.; Dinkel, J.; Ingrisch, M.; Sabel, B.O. Impact of Confounding Thoracic Tubes and Pleural Dehiscence Extent on Artificial Intelligence Pneumothorax Detection in Chest Radiographs. *Investig. Radiol.* **2020**, *55*, 792–798. [[CrossRef](#)]
79. Reyes, M.; Meier, R.; Pereira, S.; Silva, C.A.; Dahlweid, F.-M.; von Tengg-Kobligk, H.; Summers, R.M.; Wiest, R. On the Interpretability of Artificial Intelligence in Radiology: Challenges and Opportunities. *Radiol. Artif. Intell.* **2020**, *2*, e190043. [[CrossRef](#)] [[PubMed](#)]
80. Arun, N.; Gaw, N.; Singh, P.; Chang, K.; Aggarwal, M.; Chen, B.; Hoebel, K.; Gupta, S.; Patel, J.; Gidwani, M.; et al. Assessing the Trustworthiness of Saliency Maps for Localizing Abnormalities in Medical Imaging. *Radiol. Artif. Intell.* **2021**, *3*, e200267. [[CrossRef](#)] [[PubMed](#)]
81. Teng, Q.; Liu, Z.; Song, Y.; Han, K.; Lu, Y. A survey on the interpretability of deep learning in medical diagnosis. *Multimedia Syst.* **2022**, *1*–21. [[CrossRef](#)] [[PubMed](#)]
82. Rouzrokh, P.; Khosravi, B.; Faghani, S.; Moassefi, M.; Garcia, D.V.V.; Singh, Y.; Zhang, K.; Conte, G.M.; Erickson, B.J. Mitigating Bias in Radiology Machine Learning: 1. Data Handling. *Radiol. Artif. Intell.* **2022**, *4*, e210290. [[CrossRef](#)] [[PubMed](#)]
83. Ebrahimian, S.; Digumarthy, S.R.; Bizzo, B.; Primak, A.; Zimmermann, M.; Tarbiah, M.M.; Kalra, M.K.; Dreyer, K.J. Artificial Intelligence has Similar Performance to Subjective Assessment of Emphysema Severity on Chest CT. *Acad. Radiol.* **2021**, *29*, 1189–1195. [[CrossRef](#)]
84. Barbosa, E.J.M.; Gefter, W.B.; Ghesu, F.C.; Liu, S.; Mailhe, B.; Mansoor, A.; Grbic, S.; Vogt, S. Automated Detection and Quantification of COVID-19 Airspace Disease on Chest Radiographs. *Investig. Radiol.* **2021**, *56*, 471–479. [[CrossRef](#)]
85. Fogliato, R.; Chappidi, S.; Lungren, M.; Fisher, P.; Wilson, D.; Fitzke, M.; Parkinson, M.; Horvitz, E.; Inkpen, K.; Nushi, B. Who Goes First? Influences of Human-AI Workflow on Decision Making in Clinical Imaging. In Proceedings of the FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Korea, 21–24 June 2022. [[CrossRef](#)]